



Schweizerische Eidgenossenschaft  
Confédération suisse  
Confederazione Svizzera  
Confederaziun svizra

Swiss Confederation

Federal Department of Home Affairs FDHA  
Federal Office of Meteorology and Climatology MeteoSwiss



# **The GAW Station Information System (GAWSIS)**

Metadata for GAW between Discovery and  
Documentation of Observations

**Jörg Klausen**

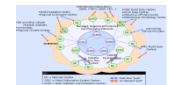
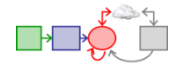
Measurement and Data Department, MeteoSwiss

GEO AQ CoP, Dublin, 5-7 Sep 2012



# Overview

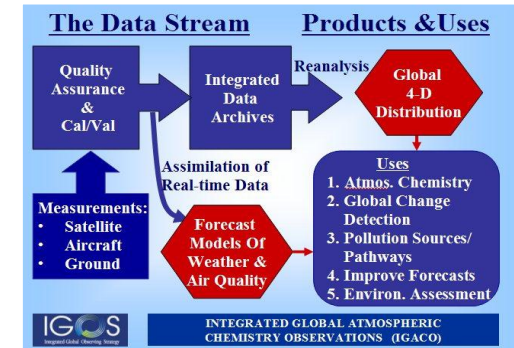
- GAW in a nutshell
- Metadata – what are we talking about?
- The „data value chain“
  - why do we care about metadata?
- WDCs and other data centres and WIS/WIGOS
- GAWSIS and metadata
- Challenges
- Conclusions





# GAW in a nutshell

- Global Atmosphere Watch (GAW) is an observing system
  - coordinated by WMO
  - guided by Scientific Advisory Groups (SAGs) and Expert Teams (ET)
- Scope of GAW is air quality as well as climate, on urban, regional and global scales.
- Mainly network(s) of ground-based stations, but also uses ship-, air-, and satellite-borne sensors.
- GAW is the atmospheric chemistry component of the Global Climate Observing System (GCOS).

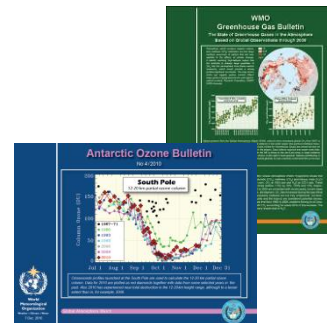
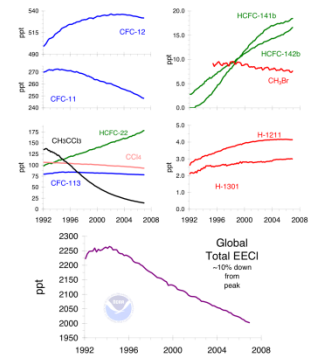
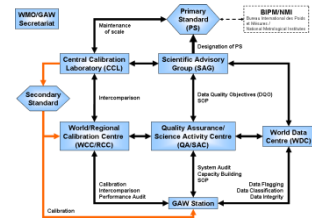




# GAW Services and Products



- Scientific stewardship for atmospheric composition monitoring
- Framework for consistent quality of observations
- Long-term observations for trend analyses
- Long-term archiving of data (→GAW WDCs)
- Ground-based data for satellite validation
- Early warning system for atmospheric change
- Contribution to quadrennial Ozone Assessment
- Contribution to IPCC Process
- Ozone, Greenhouse gas bulletins





# Metadata



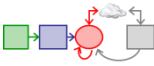
- What data?
- Where to find them?
- How to get them?
- Any strings attached?

→ **Discovery metadata**



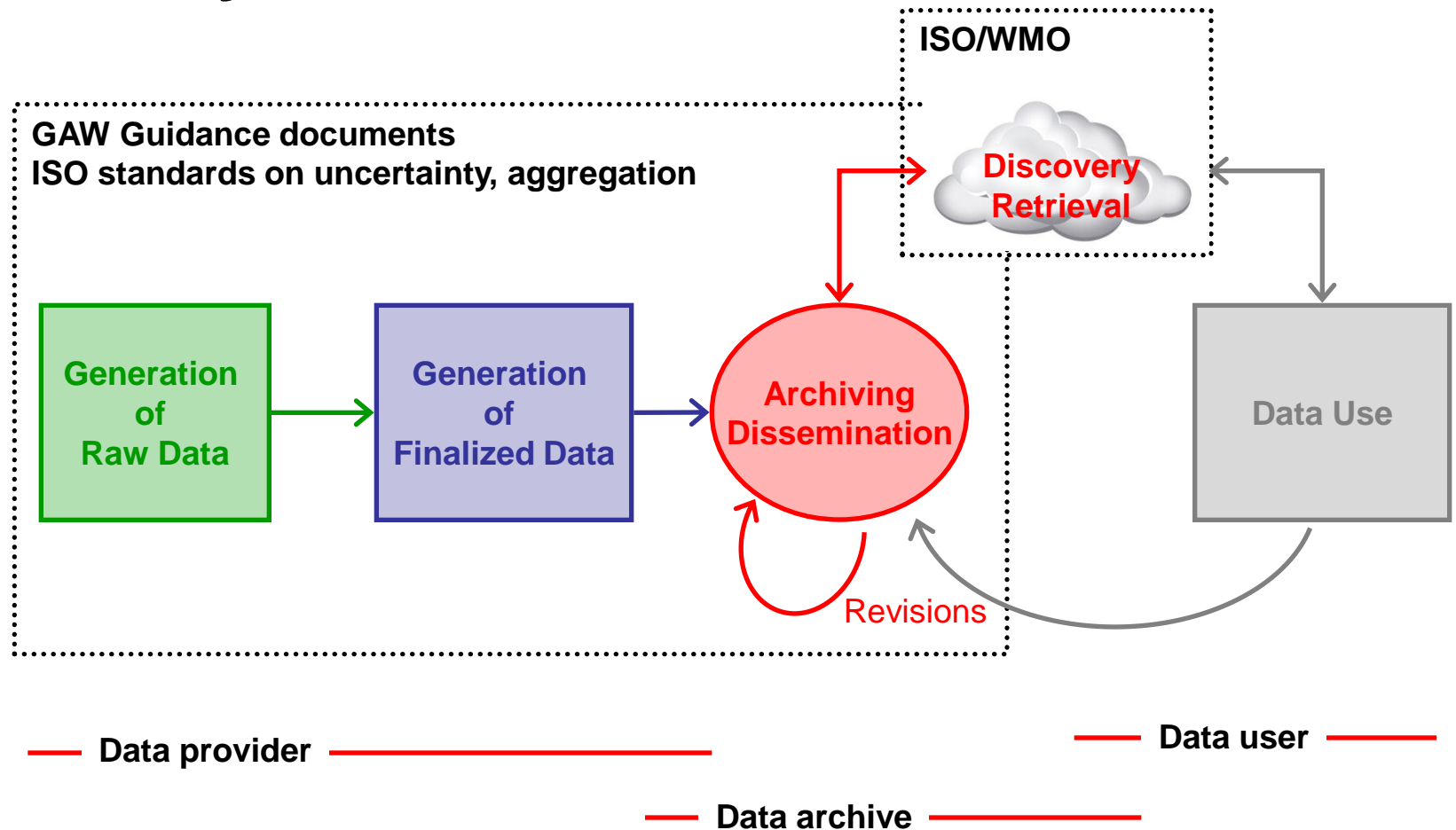
- Method of observation?
- Quality control applied?
- Data processing applied?
- Uncertainty of data?
- ...

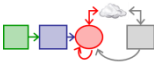
→ **Interpretation metadata**



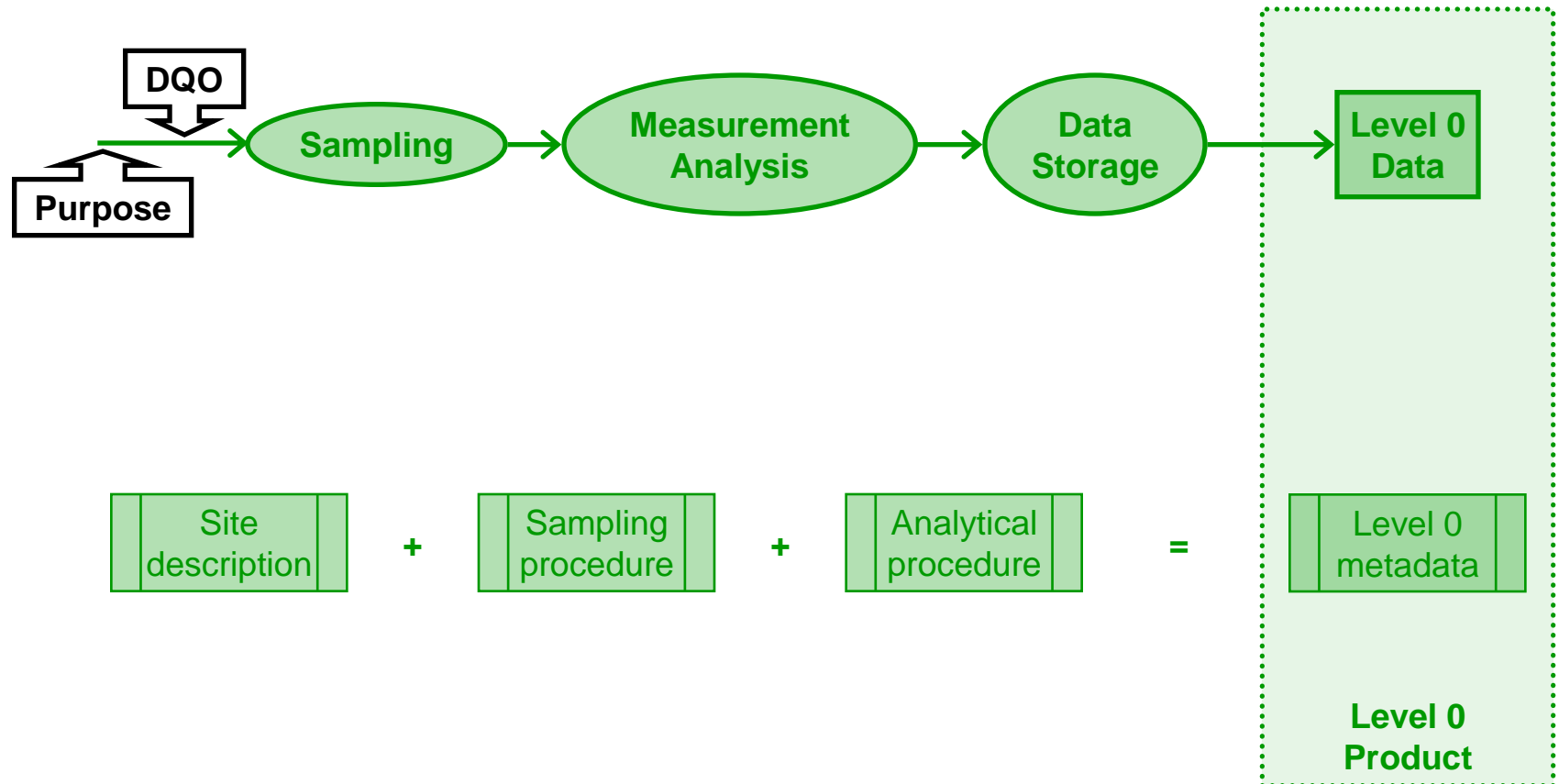
# The „data value chain“

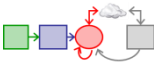
## – why do we care about metadata?



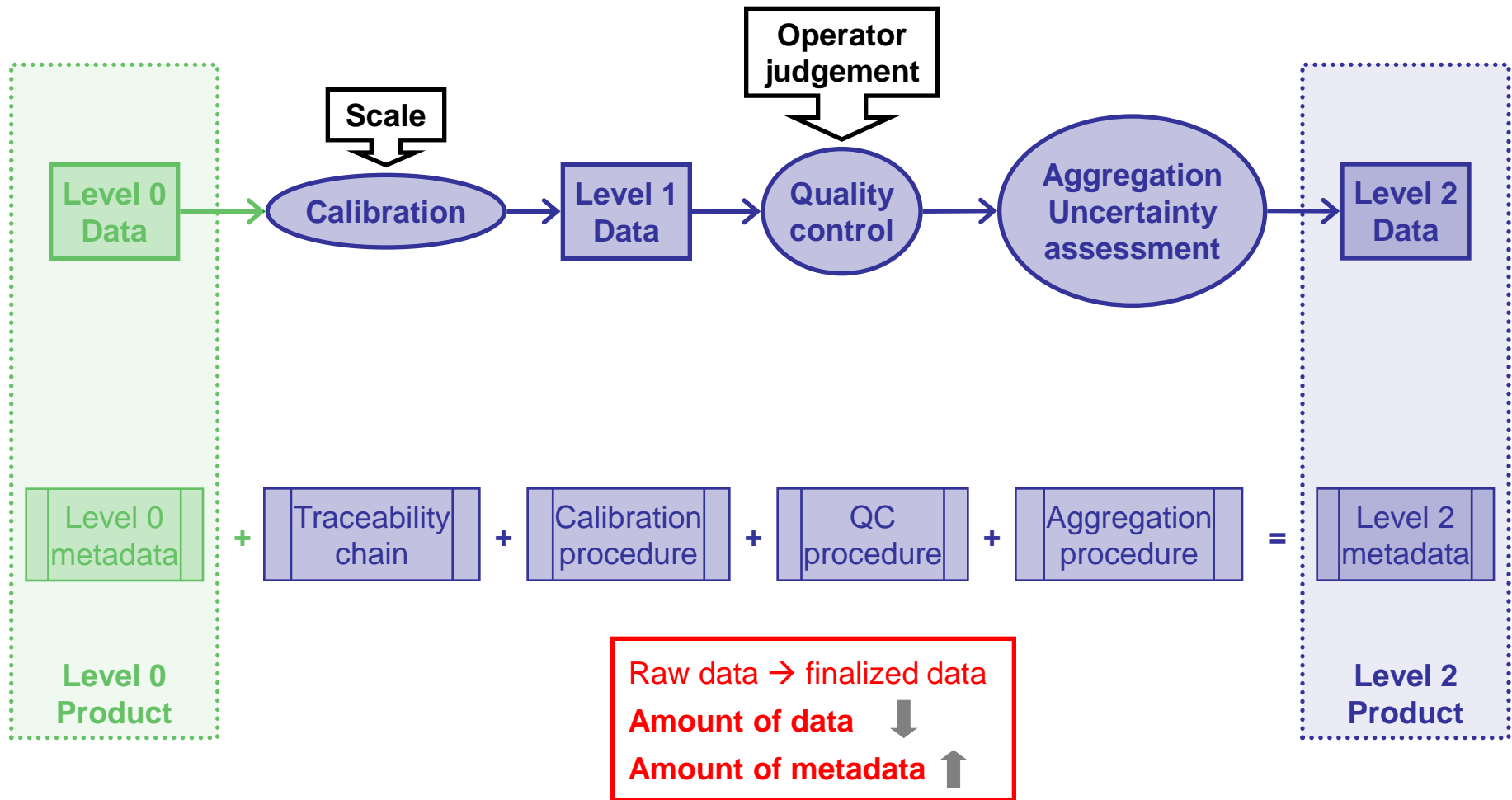


# Generation of raw data

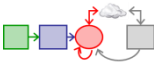




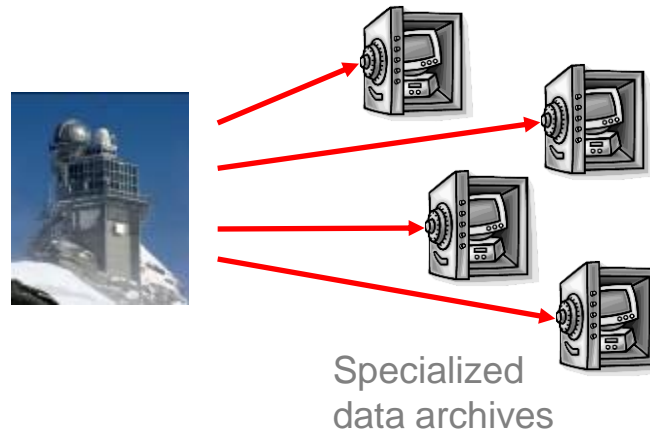
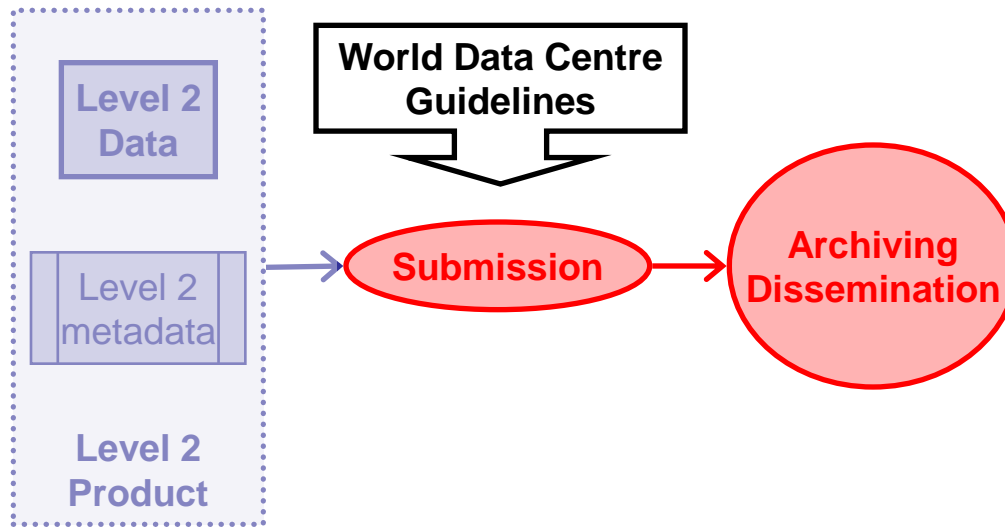
# Generation of finalized data







# Data submission



**Different data formats**  
**Different metadata requirements**  
**Different client ,handling'**



# GAW World Data Centres



## Data Policy

- For scientific purposes, access is unlimited and free.
- offer of co-authorship whenever substantial use is made of data.
- acknowledgment of data providers or owners and the data centre when used within a publication.

[GSP, 2007]

## Purpose

- collect and archive processed GAW data,
- make data publicly available,
- provide support for quality, assurance, analysis and interpretation

## Commitment

- align operations to the needs of data submitters and data users alike.

[GSP, 2007]



# Other Data Centres hosting „GAW“ Data



AERONET	NADP
AGAGE	NOAA/ESRL/GMD
BSRN	RAMCES
CapMon	SHADOZ
CDIAC	SKYNET
EANET	TCCON (CalTech)
EBAS (NILU)	[One for each satellite]
GALION (Earlinet, ...)	...

## Reasons for co-existence

History  
Specific mandate  
Data policy  
Visibility  
Limited flexibility of WDCs?  
Ad-hoc → permanent?  
Lack of coordination



## Result of co-existence

+ Expert knowledge

- Exact duplication
- Content duplication
- Different metadata
- Different versions
- Loss of traceability



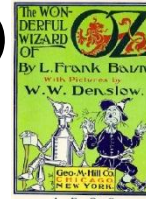
## Need for

- **ET-WDC**
- **GEO AQ CoP**
- **Discovery Metadata**
- **Standards**
- **Interoperability**





# The WISard of WIGO



Very different creatures overcome big obstacles and achieve great things by supporting each other.

## WMO Information System (WIS) succeeding Global Tele- communication System (GTS)

- Former GTS
- Exchange of n.r.t. data
- Catalogue of data
- Inter-operability of data centres

→ **Discovery metadata**

## WMO Integrated Global Observing System (WIGOS)

Framework for the integration of GOS, GAW, GCW, WHYCOS, etc.

- Coordinated governance structure
- Coordinated quality management
- Coordinated data management and delivery
- TT on WIGOS Metadata

→ **Interpretation metadata**



# Task Team on the WIGOS Metadata (TT-WMD) – Terms of Reference

- To identify the information that is needed to allow the majority of users to use WIGOS observations in appropriate contexts and in a defensible way
- To create the WIGOS Core Metadata Standard that allows the essential information to be exchanged unambiguously, regardless of the format used for the transfer
- To define the mechanism for maintaining the WIGOS Core Metadata Standard, including how metadata might be provided that is additional to the Core.
- To implement within the WIGOS Core Metadata Standard, and the WMO Core Metadata Profile, a standard method of providing users with an indication of the quality of the data, and to do so in a way that distinguishes with the quality management of the data (“quality of the observation”) and ensuring that the user is able to identify which applications the data are suitable for (“classification” of the observation).
- To complete its tasks and hand over additional requirements to its successor (if required) by 31 December 2012.

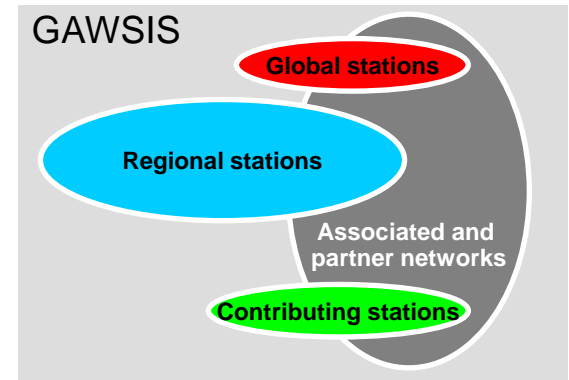
→ Interpretation metadata



# GAW SIS in a nutshell

## Some WIS functionality for GAW

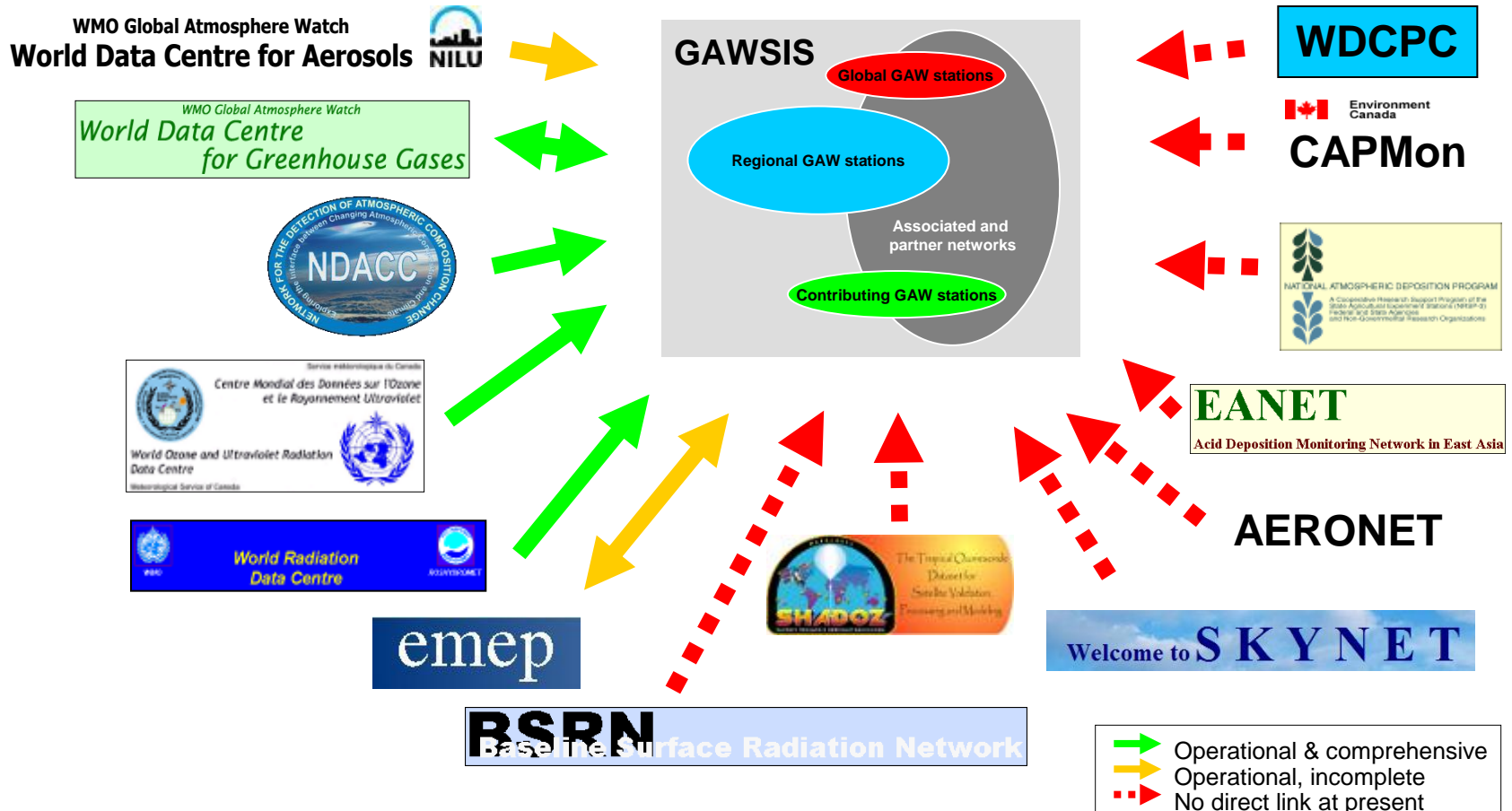
- Official catalogue of GAW ground-based network(s)
    - 800+ stations/sites
    - 9000+ data series
    - 900+ contacts
    - 600+ bibliographic references
  - Interactive, web-based metadata archive
  - Link to data across WDCs, other data centers
  - 'Clearing house' for identification of stations (colocation)
- 1-stop shop for the ground-based GAW network(s)**





# GAWSIS

## Metadata Integration for GAW





# GAWSIS Main Features

- Lists (and maps) of stations
- Lists of contacts
- Station reports
  - Site characterization
  - Measurement program
    - Meta data for each series
      - Method of observation
      - Period, frequency, etc.
      - Instrument history
      - ...
    - Hyperlinks to data (archive)
  - Contacts
  - Bibliographic references
- Clearinghouse for 3-letter station codes (incl. GAW IDs)



## Table of Contents

1. [Introduction](#)
2. [Query status of station identifiers](#)
3. [Request station identifier for a registered station](#)
4. [Register a new station and request identifier](#)
5. [Questions and comments](#)
6. [Why 3-letter codes?](#)
7. [Initial assignment of codes](#)
8. [Tables of used station identifiers](#)

## Introduction

This web site is a managed list of unique station identifiers to identify clearing house for registering such identifiers for sites not (yet) affiliated is to assign a unique 3-letter code to each known station with a long consensus as much as possible. We invite the community to participate is essential!

Please share this web address with anyone you know might have :

## Enter Station Identifier ↗

## Request Station Identifier for a Registered Station ↗

## Register a New Station and Request Identifier ↗





# GAWSIS Discovery Tools

- Simple search
- Advanced search
- Links to WDCs
- GoogleEarth™ port

## GAW World Data Centres

[WDCGG \(Gases\)](#)

[WRDC \(Radiation\)](#)

[WOUDC \(Ozone/UV\)](#)

[WDCA \(Aerosols/AOD\)](#)

[WDCPC \(Precip.  
Chem.\)](#)

[WDC-RSAT \(Remote  
Sens.\)](#)

## GoogleEarth Port



[gaw.kml](#) for a different GAWSIS experience!



# GAWSIS XML Metadata

- Profile based on ISO19115 mandatory elements
- 1 metadata set for each data series
- WIS GISC Germany compliant
- Profile presented and discussed with GISC Germany, WMO WIS PO
- Reviewed and slightly extended during ET-WDC meeting in May 2012 (WMO Geneva)
- XML representation constructed from GAWSIS relational database using 'R'

→ Mixture of discovery and interpretation metadata  
→ Still experimental and subject to further improvement  
→ Available at <http://gaw.empa.ch/gawsis/XML/>



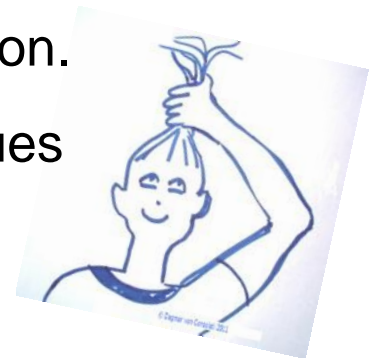
# Challenges and Open Questions

- Increasing metadata requirements
  - Legally defensible data (climate, health, ...) → traceability
  - Increased use of data in NWP/modeling → uncertainty
  - Public/open access → data quality, scope of use
- Archiving and metadata for Level 0 data → re-processing
- History/Versioning of data/metadata → long-term value of data
- Formal metadata describing data quality / scope of use?
- Who is the authority to define the metadata profile?
  - Can we agree on one profile?
- Who is the authority to define and maintain vocabularies?
  - Can CF accomodate all the needs of the GAW community?
  - What about IUPAC? ISO? WMO? EPA? EEA?



# Conclusions

- (Meta)data management starts with the definition of the purpose of the observation and the DQOs (and not the data!).
  - „Data value chain“,  
data amount ↘, but the metadata amount ↗.
  - Value of data = data quality + discoverability + documentation.
    - Feed-back of data users to archives and originators critical!
- Metadata are of the same importance as data
- ET-WDC is the WMO GAW point of contact for metadata issues related to atmospheric composition.
- Don't expect too much guidance on metadata issues  
– we're all on a learning curve!





# Thank you for your attention!



**Acknowledgment:**  
Data center managers and individual users for their support of GAWSIS